IN THE CLAIMS

1. (Original) A packet voice conferencing method comprising:

concurrently receiving a first packet voice data stream from a first conferencing endpoint and a second packet voice data stream from a second conferencing endpoint;

mapping the voice data from the first packet voice data stream to a first set of presentation mixing channels in a manner that simulates that voice data as originating in a first sector of a presentation sound field;

mapping the voice data from the second packet voice data stream to a second set of presentation mixing channels in a manner that simulates that voice data as originating in a second sector of a presentation sound field, the second sector substantially non-overlapping the first sector; and

mixing each channel from the first set of presentation mixing channels with the corresponding channel from the second set of presentation mixing channels to form a first set of mixed channels.

103 2. (Original) The method of claim 1, further comprising, for a first packet voice data stream containing information from which voice directional information can be derived:

deriving a voice arrival direction for the voice data in the first packet voice data stream;

dividing the first sector into at least two subsectors, each subsector corresponding to a range of voice arrival directions; and

when mapping the voice data from the first packet voice data stream to the first set of presentation mixing channels, performing the mapping in a manner that simulates that voice data as originating in the subsector of the presentation sound field corresponding to the voice arrival direction angle presently derived for the voice data in the first packet voice data stream.

3. (Original) The method of claim 1, further comprising:

receiving, concurrently with the first and second packet voice data streams, a third packet voice data stream from a third conferencing endpoint;

mapping the voice data from the third packet voice data stream to a third set of presentation mixing channels in a manner that simulates that voice data as originating in a

third sector of a presentation sound field, the third sector substantially non-overlapping the first and second sectors;

mixing each channel from the first set of presentation mixing channels with the corresponding channel from the third set of presentation mixing channels to form a second set of mixed channels;

mixing each channel from the second set of presentation mixing channels with the corresponding channel from the third set of presentation mixing channels to form a third set of mixed channels; and

transmitting the first, second, and third sets of mixed channels respectively to the third, second, and first conferencing endpoints.

4. (Original) The method of claim 3, further comprising establishing a control protocol with one of the first, second, and third conferencing endpoints, and accepting protocol messages from that conferencing endpoint specifying the extent of the first, second, and third sectors of the presentation sound field.

5. (Original) The method of claim 1, wherein mapping voice data to a set of presentation mixing channels comprises a method selected from the group consisting of:

splitting a voice data channel from the voice data into at least two voice data channels;

changing the relative delay of one voice data channel from the voice data with respect to another of the voice data channels;

changing the relative phase of one voice data channel from the voice data with respect to another of the voice data channels;

changing the relative amplitude of one voice data channel from the voice data with respect to another of the voice data channels;

splitting a portion of one voice data channel from the voice data and adding that portion to another of the voice data channels; and

combinations thereof.

6. (Original) The method of claim 1, further comprising pictorially displaying, on a graphical user interface, a representation of a sound field and representations of each conferencing endpoint to a listener at one conferencing endpoint, allowing that listener to manipulate the interface in order to indicate desired locations of the conferencing endpoints

within the sound field, and using the listener's manipulations to set the extent of the sectors of the presentation sound field.

7. (Original) An apparatus comprising a computer-readable medium containing computer instructions that, when executed, cause a processor or multiple communicating processors to perform a method for packet voice conferencing, the method comprising:

concurrently receiving a first packet voice data stream from a first conferencing endpoint and a second packet voice data stream from a second conferencing endpoint;

mapping the voice data from the first packet voice data stream to a first set of presentation mixing channels in a manner that simulates that voice data as originating in a first sector of a presentation sound field;

mapping the voice data from the second packet voice data stream to a second set of presentation mixing channels in a manner that simulates that voice data as originating in a second sector of a presentation sound field, the second sector substantially non-overlapping the first sector; and

mixing each channel from the first set of presentation mixing channels with the corresponding channel from the second set of presentation mixing channels to form a first set of mixed channels.

8. (Original) The apparatus of claim 7, the method further comprising:

receiving, concurrently with the first and second packet voice data streams, a third packet voice data stream from a third conferencing endpoint;

mapping the voice data from the third packet voice data stream to a third set of presentation mixing channels in a manner that simulates that voice data as originating in a third sector of a presentation sound field, the third sector substantially non-overlapping the first and second sectors;

mixing each channel from the first set of presentation mixing channels with the corresponding channel from the third set of presentation mixing channels to form a second set of mixed channels;

mixing each channel from the second set of presentation mixing channels with the corresponding channel from the third set of presentation mixing channels to form a third set of mixed channels; and

transmitting the first, second, and third sets of mixed channels respectively to the

third, second, and first conferencing endpoints.

9. (Original) The apparatus of claim 8, the method further comprising establishing a control protocol with one of the first, second, and third conferencing endpoints, and accepting protocol messages from that conferencing endpoint specifying the extent of the first, second, and third sectors of the presentation sound field.

10. (Original) The apparatus of claim 8, the method further comprising establishing a control protocol session with a user interface for a participant located at one of the conferencing endpoints, and accepting protocol messages from that user interface specifying the division of the presentation sound field for that endpoint.

11. (Original) The apparatus of claim 10, the method further comprising establishing a control protocol session with a user interface for a participant located at each of the other conferencing endpoints, thereby allowing each endpoint to specify its own division of the presentation sound field.

12. (Original) The apparatus of claim 7, wherein mapping voice data to a set of presentation mixing channels comprises a method selected from the group consisting of:

splitting a voice data channel from the voice data into at least two voice data channels;

changing the relative delay of one voice data channel from the voice data with respect to another of the voice data channels;

changing the relative phase of one voice data channel from the voice data with respect to another of the voice data channels;

changing the relative amplitude of one voice data channel from the voice data with respect to another of the voice data channels;

splitting a portion of one voice data channel from the voice data and adding that portion to another of the voice data channels; and

combinations thereof.

13. (Original) The apparatus of claim 12, wherein a mapping is performed on a subchannel basis.

14. (Original) The apparatus of claim 7, the method further comprising, when voice data from one of the conferencing endpoints is received monaurally, mapping the voice data into multiple voice data channels.

15. (Original) The apparatus of claim 7, the method further comprising, when voice data from one of the conferencing endpoints comprises multiple voice data channels:

measuring the relative delay between at least two of the multiple channels;

estimating, from the measured relative delay, the arrival direction of a voice signal present in the voice data; and

accounting for the estimated arrival direction during mapping of the voice data into a set of presentation mixing channels.

16. (Original) The apparatus of claim 7, the method further comprising, for a first packet voice data stream containing information from which voice directional information can be derived:

deriving a voice arrival direction for the voice data in the first packet voice data stream;

dividing the first sector into at least two subsectors, each subsector corresponding to a range of voice arrival directions; and

when mapping the voice data from the first packet voice data stream to the first set of presentation mixing channels, performing the mapping in a manner that simulates that voice data as originating in the subsector of the presentation sound field corresponding to the voice arrival direction angle presently derived for the voice data in the first packet voice data stream.

17. (Original) The apparatus of claim 7, the method further comprising pictorially displaying, on a graphical user interface, a representation of a sound field and representations of each conferencing endpoint to a listener at one conferencing endpoint, allowing that listener to manipulate the interface in order to indicate desired locations of the conferencing endpoints within the sound field, and using the listener's manipulations to set the extent of the sectors of the presentation sound field.

18. (Original) The apparatus of claim 17, the method further comprising allowing the listener to divide a sector into subsectors, and to manipulate each subsector of that sector within the

presentation sound field independent of the other subsectors of that sector.

19. (Original) The apparatus of claim 17, wherein the graphical user interface further allows the listener to specify the number and locations of presentation channel acoustical speakers relative to that listener's position in a room, the method further comprising accounting for the number and locations of presentation channel acoustical speakers in mapping voice data to presentation mixing channels.

20. (Original) The apparatus of claim 17, the method further comprising recurrently updating the graphical user interface with a visual indication of which endpoint or endpoints is/are currently transmitting voice data.

21. (Original) The apparatus of claim 17, the method further comprising automatically dividing the presentation sound field into sectors that allocate approximately equal shares of the presentation sound field to each endpoint.

22. (Original) The apparatus of claim 21, the method further comprising tracking the number of conferencing endpoints participating in a conference, and automatically altering the allocation of the presentation sound field as endpoints are added to or leave the conference.

23. (Original) The apparatus of claim 21, wherein a larger sector of the sound field is allocated to a conferencing endpoint that is broadcasting multiple capture channels than is allocated to a conferencing endpoint that is broadcasting monaurally.

24. (Original) A packet voice conferencing system comprising:

means for concurrently receiving multiple packet voice data streams;

means for manipulating the voice data in each of the packet voice data streams in a manner that simulates that voice data as originating in a specified sector of a presentation sound field, the sectors arranged in the sound field in substantially non-overlapping fashion; and

means for combining the manipulated voice data from each packet voice data stream into a set of presentation channels.

25. (Original) The packet voice conferencing system of claim 24, further comprising means for specifying the sector of the presentation sound field to be applied to each packet voice data stream.

26. (Currently amended) The packet voice conferencing system of claim 24, further comprising means for varying the specified sector of the presentation sound field for a packet voice data stream depending on a voice arrival direction derived for that packet voice data stream.

27. (Original) The packet voice conferencing system of claim 24, incorporated into one of the conferencing endpoints.

28. (Original) A packet voice conferencing system comprising:

first and second decoders, to respectively decode first and second packet voice data streams and produce first and second sets of one or more voice data channels from the voice data packets contained in the streams;

a packet switch to receive packet voice data streams sent to the system by first and second conferencing endpoints and to distribute the packet voice data stream received from the first conferencing endpoint to the first decoder and the packet voice data stream received from the second conferencing endpoint to the second decoder;

a first channel mapper to map the first set of voice data channels to a first set of presentation mixing channels in a manner that simulates the voice data as originating in a first sector of a presentation sound field;

a second channel mapper to map the second set of voice data channels to a second set of presentation mixing channels in a manner that simulates the voice data as originating in a second sector of a presentation sound field, the second sector substantially non-overlapping the first sector; and

a first set of mixers, each mixer combining one of the first set of presentation mixing channels with a corresponding one of the second set of presentation mixing channels to form a mixed channel, the set of mixers collectively forming a first set of mixed channels.

29. (Original) The packet voice conferencing system of claim 28, further comprising:

a third decoder to decode a third packet voice data stream and produce a third set of one or more voice data channels from the voice data packets contained in the third stream, the

packet switch receiving the third packet voice data stream from a third conferencing endpoint and distributing the third packet voice data stream to the third decoder;

a third channel mapper to map the third set of voice data channels to a third set of presentation mixing channels in a manner that simulates the voice data as originating in a third sector of a presentation sound field, the third sector substantially non-overlapping the first and second sectors;

a second set of mixers, each mixer in the second set combining one of the first set of presentation mixing channels with a corresponding one of the third set of presentation mixing channels to form a mixed channel, the second set of mixers collectively forming a second set of mixed channels;

a third set of mixers, each mixer in the third set combining one of the second set of presentation mixing channels with a corresponding one of the third set of presentation mixing channels to form a mixed channel, the third set of mixers collectively forming a third set of mixed channels; and

a transmitter to dispatch the first, second, and third sets of mixed channels respectively to the third, second, and first conferencing endpoints.

30. (Original) The packet voice conferencing system of claim 29, further comprising a controller connected to each channel mapper, the controller configuring each channel mapper according to its designated sound field sector.

31. (Original) The packet voice conferencing system of claim 30, wherein the controller communicates with one of the first, second, and third conferencing endpoints using a control protocol and accepts protocol messages from that conferencing endpoint specifying the extent of the first, second, and third sectors of the presentation sound field.

32. (Original) The packet voice conferencing system of claim 28, further comprising a jitter buffer for each voice data channel, each jitter buffer delaying its respective voice data channel prior to submission to a mapper.

33. (Original) The packet voice conferencing system of claim 32, further comprising a controller connected to the jitter buffers to synchronize the relative delays of multiple jitter buffers associated with a common mixed channel.

34. (Original) The packet voice conferencing system of claim 28, further comprising a graphical user interface driver to create a display for a listener and manipulate that display in response to listener inputs, the display including a representation of a sound field and representations of each conferencing endpoint, the driver using listener inputs to set the extent of the sectors of the presentation sound field.

35. (Original) The packet voice conferencing system of claim 34, wherein the graphical user interface driver allows the listener to divide a sector into subsectors, and to manipulate each subsector of that sector within the presentation sound field independent of the other subsectors of that sector.

36. (Currently amended) A packet voice conferencing system comprising:

a decoder, to decode a packet voice data stream to produce a set of one or more voice data channels from the voice data packets contained in the ~~streams~~ stream and a voice arrival direction corresponding to the set of voice data channels;

a controller to select one of a plurality of presentation sound field subsectors for the voice data channels based on the voice arrival direction, each subsector corresponding to a range of voice arrival directions; and

a channel mapper to map the set of voice data channels to a set of presentation channels in a manner that simulates the voice data as originating in the selected subsector of the presentation sound field.

37. (Original) The packet voice conferencing system of claim 36, wherein the voice arrival direction is explicitly communicated in the packet voice data stream.

38. (Original) The packet voice conferencing system of claim 36, wherein the set of voice data channels comprises two or more channels, and wherein the decoder comprises a direction finder to estimate the voice arrival direction by comparing at least one of the voice data channels to another of the voice data channels.

39. (Original) A packet voice conferencing system having one or more local audio capture channels, the system comprising:

a controller to negotiate with other packet voice conferencing systems connected in a common conference, wherein the results of a negotiation include a codec to be used by the

system for encoding the local audio capture channels, and a presentation sound field sector allocated to the local audio capture channels;

a channel mapper to map the local audio capture channels to a set of presentation mixing channels in a manner that simulates the audio data on the capture channels as originating in the allocated presentation sound field sector; and

an encoder to encode the presentation mixing channels into a packet voice data stream.